MAA704: Matrix factorization and canonical forms

Christopher Engström

November 14, 2014

Contents of todays lecture

- Some interesting / useful / important properties of matrices
- Matrix factorization
- Canonical forms

Matrix factorization

- Rewriting a matrix as a product of several matrices.
- Choosing these factor matrices wisely can make problems easier to solve.
- Also known as *matrix decomposition*

Diagonalizable matrix

Definition

If $\mathsf{B} = \mathsf{S}^{-1}\mathsf{DS}$ where D is a diagonal matrix then B is diagonalizable.

Motivation.

Using elementary row operations we want to turn $\mathsf{B}\mathbf{x} = \mathbf{y}$ into $\mathsf{D}\hat{\mathbf{x}} = \hat{\mathbf{y}}$. This can be written as $\mathsf{SB}\mathbf{x} = \mathsf{S}\mathbf{y}$. Since elementary row operations are invertible $\mathsf{SBS}^{-1}\mathsf{S}\mathbf{x} = \mathsf{S}\mathbf{y}$. Let $\hat{\mathbf{x}} = \mathsf{S}\mathbf{x}$ and $\hat{\mathbf{y}} = \mathsf{S}\mathbf{y}$, then

$$\mathsf{D} = \mathsf{SBS}^{-1} \Leftrightarrow \mathsf{B} = \mathsf{S}^{-1}\mathsf{DS}$$

□

Triangular matrix

$$A = \begin{bmatrix} \bigstar & \bigstar & \dots & \bigstar \\ 0 & \bigstar & \dots & \bigstar \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \bigstar \end{bmatrix}$$

Triangular matrix

- Can be lower (left) or upper (right) triangular
- Easy to solve equation systems involving triangular matrices
- Diagonal values are also eigenvalues

Hessenberg matrix

$$A = \begin{bmatrix} \bigstar & \bigstar & \bigstar & \cdots & \bigstar & \bigstar & \bigstar \\ \bigstar & \bigstar & \bigstar & \cdots & \bigstar & \bigstar & \bigstar \\ 0 & \bigstar & \bigstar & \cdots & \bigstar & \bigstar & \bigstar \\ 0 & 0 & \bigstar & \cdots & \bigstar & \bigstar & \bigstar \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \bigstar & \bigstar & \bigstar \\ 0 & 0 & 0 & \cdots & 0 & \bigstar & \bigstar \end{bmatrix}$$

Hessenberg matrix

- 'Almost' triangular
- Multiplication of a (upper) Hessenberg matrices and a (upper) triangular matrix gives a new Hessenberg matrix (Useful in for example the QR-method used to find eigenvalues of a matrix).
- Diagonal elements usually give a rough approximation of the eigenvalues.

Hermitian matrix

Definition

The *Hermitian conjugate* of a matrix A is denoted A^H and is defined by $(A^H)_{ij} = \overline{(A)_{ji}}$.

Definition

A matrix is said to be *Hermitian* (or *self-adjoint*) if $A^H = A$

Hermitian matrix

- Notice the similarities with a symmetric matrix $A^\top = A$.
- All eigenvalues real.
- Always diagonalizable.
- Important in theoretical physics, quantum physics, electroengineering and in certain problems in statistics.

Unitary matrices

Definition

A matrix, A, is said to be *unitary* if $A^H = A^{-1}$.

Properties of unitary matrices

Theorem

Let U *be a unitary matrix, then*

a) U is always invertible.
b) U^{-1} is also unitary.
c) $|\det(U)| = 1$
d) $(UV)^H = (UV)^{-1}$ if V is also unitary.
e) For any λ that is an eigenvalue of U, $\lambda = e^{i\omega}$, $0 \leq \omega \leq 2\pi$.
f) Let $\mathbf{v}$ be a vector, then $|U\mathbf{v}| = |\mathbf{v}|$ (for any vector norm).
g) The rows/columns of U are *orthonormal*, that is $U_{i.}U_{j.}^H = 0$, $i \neq j$, $U_{k.}U_{k.}^H = 1$.
h) U preserves eigenvalues.

Example of a unitary matrix

- The C matrix below rotates a vector by the angle θ around the x-axis

$$C = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(\theta) & -\sin(\theta) \\ 0 & \sin(\theta) & \cos(\theta) \end{bmatrix}$$

and is a unitary matrix.

Positive definite matrix

Definition

We consider a square symmetric real valued $n \times n$ matrix A, then:

- A is positive definite if $\mathbf{x}^\top A\mathbf{x}$ is positive for all non-zero vectors $\mathbf{x}$.
- A is positive semidefinite if $\mathbf{x}^\top A\mathbf{x}$ is non-negative for all non-zero vectors $\mathbf{x}$.

- A is *positive definite* $\Leftrightarrow \lambda > 0$ for all λ eigenvalue of A.
- Can also define negative definite and semi-definite matrices.

Positive definite matrix

Positive definite matrices have many useful properties, if A is positive definite then

- A is invertible.
- A have a unique cholesky decomposition (seen later today).
- Positive definite matrices are closely related to quadratic forms (last lecture).
- Any Covariance matrix is positive semi-definite.

Matrix factorization

- Diagonalizable $A = S^{-1}DS$ with D diagonal
- Other important factorizations:
  - Spectral factorization $Q\Lambda Q^{-1}$
  - LU-factorization
  - Cholesky factorization GG^H
  - QR-factorization
  - Rank factorization CF
  - Jordan canonical form $S^{-1}JS$
  - Singular value factorization $U\Sigma V^H$

Spectral factorization

- Spectral factorization is a special version of diagonal factorization.
- It is sometimes referred to as *eigendecomposition*.
- Let A be an square $(n \times n)$ matrix with linearly independent rows. Then

$$A = Q\Lambda Q^{-1}$$

where $AQ_{.i} = \Lambda_{ii}Q_{.i}$ for all $1 \leq i \leq n$.

Rank factorization

- Let A be an $m \times n$ matrix with $\text{rank}(A) = r$ (A has r independent rows/columns). Then

$$A = CF$$

where $C \in \mathcal{M}_{m \times r}$ and $F \in \mathcal{M}_{r \times n}$

Rank factorization

- How can we find this factorization?
- Rewrite matrix on reduced row echelon form

$$B = \begin{bmatrix} 0 & 1 & * & 0 & 0 & 0 & * & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & * & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & * & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & * & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Rank factorization

- Create C by removing all columns in A that corresponds to a non-pivot column in B.
- In this example

$$C = \begin{bmatrix} A_{.2} & A_{.4} & A_{.5} & A_{.6} & A_{.8} \end{bmatrix}$$

- Create F by removing all zero rows in B.
- In this example

$$F = \begin{bmatrix} B_{1.} & B_{2.} & B_{3.} & B_{4.} & B_{5.} \end{bmatrix}$$

LU-factorization

- $A = LR = LU$.
- L is a $n \times n$ lower triangular matrix.
- U is a $n \times m$ upper triangular matrix.
- Solve $Ax = L(Ux) = b$ by first solving $Ly = b$ and then solve $Ux = y$. Both these systems are easy to solve since L and U are both triangular.
- Not every matrix A have a LU factorization, not even every square invertible matrix.

LUP-factorization

Theorem

Every $n \times m$ *matrix* A *have a matrix factorization*

$$\mathrm{PA} = \mathrm{LU}$$

. *where*

- P *is a* $n \times n$ *permutation matrix.*
- L *is a* $n \times n$ *lower triangular matrix.*
- U *is a* $n \times m$ *upper triangular matrix.*

Cholesky factorization

- Systems involving triangular matrices are often easy to solve.
- Try to rewrite a matrix as a product that contains a triangular matrix seems like a good idea.
- One way is using *LU*-factorization where $\mathrm{PA} = \mathrm{LU}$ where P is a permutation matrix, L is a lower- and U is an upper triangular matrix.
- There is also the Cholesky factorization, $\mathrm{A} = \mathrm{GG}^H$, where A is *Hermitian* and *positive-definite* and G is lower triangular.

Cholesky factorization

- Consider the equation $Ax = y$. If a can be Cholesky factorized, $A = GG^H$, this equation can be turned into two new equations:

$$\begin{cases} Gz = y \\ G^H x = z \end{cases}$$

both of these equations are easy to solve.

Calculating the Cholesky factorization

Looking at the relation $A = LL^\top$ for a real positive definite 3×3 matrix we get:

$$A = \begin{bmatrix} L_{1,1} & 0 & 0 \\ L_{2,1} & L_{2,2} & 0 \\ L_{3,1} & L_{3,2} & L_{3,3} \end{bmatrix} \begin{bmatrix} L_{1,1} & L_{2,1} & L_{3,1} \\ 0 & L_{2,2} & L_{3,2} \\ 0 & 0 & L_{3,3} \end{bmatrix}$$

$$= \begin{bmatrix} L_{1,1}^2 & L_{2,1}L_{1,1} & L_{3,1}L_{1,1} \\ L_{2,1}L_{1,1} & L_{2,1}^2 + L_{2,2}^2 & L_{3,1}L_{2,1} + L_{3,2}L_{2,2} \\ L_{3,1}L_{1,1} & L_{3,1}L_{2,1} + L_{3,2}L_{2,2} & L_{3,1}^2 + L_{3,2}^2 + L_{3,3}^2 \end{bmatrix}$$

Calculating the Cholesky factorization

Since A is symmetric we only need to calculate the lower triangular part.

$$\begin{bmatrix} L_{1,1}^2 & - & - \\ L_{2,1}L_{1,1} & L_{2,1}^2 + L_{2,2}^2 & - \\ L_{3,1}L_{1,1} & L_{3,1}L_{2,1} + L_{3,2}L_{2,2} & L_{3,1}^2 + L_{3,2}^2 + L_{3,3}^2 \end{bmatrix}$$

- For the elements $L_{i,j}$ we get:

$$L_{j,j} = \sqrt{A_{j,j} - \sum_{k=1}^{j-1} L_{j,k}^2}$$

$$L_{i,j} = \frac{1}{L_{j,j}} \left(A_{i,j} - \sum_{k=1}^{j-1} L_{i,k}L_{j,k} \right), \ i > j$$

- We notice that we only need the elements above and to the left to calculate the next element.

Applications of Cholesky factorization

- Are there any interesting matrices that can be easy Cholesky factorized?
- Any covariance matrix is positive-definite and any covariance matrix based on measured data is going to be symmetric and real-valued. From the last two properties it follows that this matrix is Hermitian.
- Example application: generating variates according to a multivariate distribution with covariance matrix Σ and expected value μ
  Using the Cholesky factorization you get the simple formula $X = \mu + G^{\top} Z$ where X is the variate, $\Sigma = GG^{H}$ and Z is a vector of standard normal variates.

QR factorization

Theorem

Every $n \times m$ *matrix* A *have a matrix decomposition*

$$A = QR$$

where

- R *is a* $n \times m$ *upper triangular matrix..*
- Q *is a* $n \times n$ *unitary matrix.*

QR factorization

- Given a QR-factorization we can solve a linear system $Ax = b$ by solving $Rx = Q^{-1}b = Q^{H}b$. Which is can be done fast since R is a triangular matrix.
- QR-factorization can also used in solving the linear least square problem.
- It plays an important role in the QR-method used to calculate eigenvalues of a matrix numerically.

Canonical form

- A *canonical form* is a standard way of describing an object.
- There can be several different kinds of canonical forms for an object.
- Some examples for matrices:
  - Diagonal form (for diagonalizable matrices)
  - Reduced row echelon form (for all matrices)
  - Jordan canonical form (for square matrices)
  - Singular value factorization form (for all matrices)

Reduced row echelon form

Definition

A matrix is written on *reduced row echelon form* when they are written on echelon form and their pivot elements are all equal to one and all other elements in a pivot column are zero.

$$B = \begin{bmatrix} 0 & 1 & * & 0 & 0 & 0 & * & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & * & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & * & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & * & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Theorem

All matrices are similar to some reduced row echelon matrix.

Jordan normal form

Definition (Jordan block)

A *Jordan block* is a square matrix of the form

$$J_m(\lambda) = \begin{bmatrix} \lambda & 1 & 0 & \dots & 0 \\ 0 & \lambda & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda & 1 \\ 0 & 0 & \dots & 0 & \lambda \end{bmatrix}$$

Jordan normal form

Definition (Jordan matrix)

A *Jordan matrix* is a square matrix of the form

$$J = \begin{bmatrix} J_{m_1}(\lambda_1) & 0 & \dots & 0 \\ 0 & J_{m_2}(\lambda_2) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & J_{m_1}(\lambda_k) \end{bmatrix}$$

Jordan normal form

Theorem

All square matrices are similar to a Jordan matrix. The Jordan matrix is unique except for the order of the Jordan blocks. This Jordan matrix is called the Jordan normal form of the matrix.

Theorem (Some other interesting properties of the Jordan normal form)

Let $A = S^{-1}JS$

a) *The eigenvalues of* J *is the same as the diagonal elements of* J*.*
b) J *has one eigenvector per Jordan block.*
c) *The rank of* J *is equal to the number of Jordan blocks.*
d) *The normal form is sensitive to perturbations. This means that a small change in the normal form can mean a large change in the* A *matrix and vice versa.*

Singular value factorization

Theorem

All $\mathsf{A} \in \mathcal{M}_{m\times n}$ *can be factorized as*

$$\mathsf{A} = \mathsf{U}\Sigma\mathsf{V}^H$$

where U *and* V *are unitary matrices and*

$$\Sigma = \begin{bmatrix} \mathsf{S}_r & 0 \\ 0 & 0 \end{bmatrix}$$

where S_r *is a diagonal matrix with* $r = rank(\mathsf{A})$. *The diagonal elements of* S_r *are called the singular values. The singular values are uniquely determined by the matrix* A *(but not necessarily their order).*

Singular value factorization

- Very often referred to as the SVD (singular value decomposition).
- Used a lot in statistics and information processing.
- Can be used to quantify many different qualities of matrices, more on this in later lectures.

Similar matrices

- In everyday language two matrices are 'similar' if they have almost the same elements or structure. But there is also a precise mathematical relation between two matrices that is called similar.

Definition

Two matrices, A and B, are *similar* if $A = S^{-1}BS$.

Interesting properties of similar matrices

- Similar matrices share several properties:
  - Eigenvalues (but generally not eigenvectors)
  - Determinant
  - Trace
  - Rank
- We have already seen some examples of why similar matrices are interesting:
  - Diagonalizable matrices $A = S^{-1}BS$
  - Permutation matrices $A = PBP^{\top}$
  - Jordan normal form $A = S^{-1}JS$
- Similarity between matrices mean they represent the same linear mapping described in different basis.

Summary

- Triangular and Hessenberg matrices
- Hermitian matrices
- Unitary matrices
- Positive definite matrices

Summary

- Matrix factorization
  - Spectral factorization $Q\Lambda Q^{-1}$
  - LU-factorization
  - Cholesky factorization GG^H
  - QR-factorization
  - Rank factorization CF
  - Jordan canonical form $S^{-1}JS$
  - Singular value factorization $U\Sigma V^H$